

ное разложение гиперграфа множествами сочленения. Доказательство теорем 2 и 3 служит теоретическим обоснованием этих методов. Предложенные методы имеют полиномиальную вычислительную сложность и могут применяться независимо друг от друга. Однако для достижения наибольшего эффекта предобработки рекомендуется сочетать

эти методы. Конечно, существуют гиперграфы, для которых безопасное редуцирование и безопасное разложение гиперграфа множествами сочленения не приводят к желаемому эффекту. Поэтому перспективны исследования, направленные на разработку других методов безопасной предобработки гиперграфов относительно древовидной ширины.

#### СПИСОК ЛИТЕРАТУРЫ

1. Bodlaender H.L. Discovering treewidth // Proc. of the 31<sup>st</sup> Conf. SOFSEM 2005, Springer-Verlag, Lecture Notes in Computer Science 3381. – 2005. – P. 1–16.
2. Robertson N., Seymour P.D. Graph minors. II. Algorithmic aspects of treewidth // J. Algorithms. – 1986. – V. 7. – P. 309–322.
3. Gogate V., Dechter R. A complete anytime algorithm for treewidth // Proc. of the 20<sup>th</sup> Conf. on Uncertainty in Artificial Intelligence. – 2004. – P. 201–204.
4. Мейер Д. Теория реляционных баз данных. – М.: Мир, 1987. – 608 с.
5. Емеличев В.А., Мельников О.И., Сарванов В.И., Тышкевич Р.И. Лекции по теории графов. – М.: Наука, 1990. – 384 с.
6. Зыков А.А. Гиперграфы // Успехи математических наук. – 1974. – Т. 29. – Вып. 6. – С. 89–154.
7. Быкова В.В. М-ациклические и древовидные гиперграфы // Известия Томского политехнического университета. – 2010. – Т. 317. – № 2. – С. 25–30.

Поступила 24.02.2011 г.

УДК 004.94

## МЕТОД ИНТЕГРАЦИИ СТРУКТУРНЫХ РАЗЛИЧИЙ В ГРАФОВЫХ МОДЕЛЯХ И ЕГО ПРИМЕНЕНИЕ ДЛЯ ОПИСАНИЯ СТРУКТУР

В.К. Погребной

Томский политехнический университет

E-mail: [vkp@tpu.ru](mailto:vkp@tpu.ru)

Рассматривается проблема анализа структур графов на предмет оценки их подобия и выявления труднообнаруживаемых свойств. Предлагаются методы свободной и зависимой интеграции структурных различий и получения интегральных описателей структур (ISD). Введено понятие области интеграции различий для класса структур и разработаны правила её формирования. Предложенные методы составили основу новой технологии анализа структур с помощью ISD, названной ISD-технологией.

#### Ключевые слова:

Структура, структурные различия, метод интеграции структурных различий, интегральный описатель структур (ISD), область интеграции для класса структур, ISD-технология анализа структур

#### Key words:

Structure, structure differences, the method of integration of structural differences, integral structure descriptor (ISD), area of integration for the class of structures, ISD-technology analysis of structures

#### Введение

Графовые модели широко используются для визуального и математического представления структурных свойств объектов в самых разных предметных областях: радиоэлектроника, энергетика, транспорт, связь, химия, биология. Модели, отражающие сетевую природу исследуемых объектов, часто оказываются конструктивными и наряду с решением задач анализа могут непосредственно использоваться в задачах синтеза.

Построение графовой модели предполагает разбиение объекта на компоненты (вершины) и установление отношений между ними (рёбра). Вершинам и рёбрам могут приписываться различные атрибуты, которые отражают конкретные свойства соответствующих компонентов и связей между ними.

Использование таких атрибутов при построении графовых моделей позволяет учесть специфику объектов, но при этом приходится иметь дело с различными видами графов и гиперграфов.

Учитывая, что данная статья ограничивается рассмотрением только структурных свойств объектов, необходимость в использовании атрибутов отпадает. Поэтому, следуя базовому определению понятия структуры как совокупности элементов множества и связей между ними, в работе в качестве графовой модели принимается обыкновенный граф, который по определению отражает понятие структуры и соответствует цели исследования — разработке конструктивной математической модели формального описания структур, позволяющей однозначно представлять структуры и сравнивать

их между собой. Для упрощения столь сложной задачи в последующем будем рассматривать структуры графов в составе априорно задаваемого класса.

#### Содержательная формулировка задачи

Пусть на коммутационном поле размещена совокупность контактов схемы и указаны связи между ними. Данная схема соответствует обыкновенному графу и может быть представлена, например, в виде матрицы смежности вершин. Предположим, что описание схемы выполняется двумя разработчиками независимо друг от друга. Один из них решил пронумеровать контакты схемы слева направо и сверху вниз. Полученный граф и соответствующая матрица показаны на рис. 1, а. Второй разработчик пронумеровал контакты сверху вниз и слева направо. Соответствующий граф и матрица представлены на рис. 1, б.

Очевидны недостатки представления (описания) структуры схемы матрицей смежности вершин соответствующего графа. Среди них выделим следующие:

- зависимость описания от нумерации вершин и визуального представления графа;
- по описанию двух структур нельзя судить об их идентичности или наличии общих частей (подструктур);
- описание структуры совокупностью бинарных отношений ограничивает возможности выявления труднообнаруживаемых и, как правило, особо важных структурных свойств объектов.

Первый недостаток хорошо иллюстрируется на рис. 1. Здесь показана зависимость описания от нумерации вершин при условии сохранения визуального представления графов. Если в графах изменить топологию расположения вершин, то полу-

чим графы, которые и визуально будут трудносопоставимы. Заметим, что данный недостаток не возникает, если в схеме вместо контактов разместить разные элементы так, чтобы каждой вершине графа соответствовал элемент определённого типа. В этом случае для сравнения структур достаточно в матрицах одинаково упорядочить строки и столбцы по типам элементов. Очевидно, что полученные таким образом матрицы совпадут. Этот пример наводит на мысль о продуктивности идеи получения для каждой вершины графа уникальной структурной характеристики. Наличие таких характеристик, назовём их *описателями вершин*, позволило бы существенно продвинуться в решении задач описания и анализа структур.

Второй недостаток преодолевается решением задач определения изоморфизма и изоморфного вхождения графов. Обе задачи относятся к классу неполиномиально сложных. Вместе с тем для их решения в [1, 2] предложены алгоритмы, приемлемые для практического применения. Алгоритмы нацелены на получение однозначного ответа по данным задачам и, следовательно, не совсем соответствуют рассматриваемым целям анализа структур, когда, в частности, требуются промежуточные ответы о степени подобия структур или наличия в них отдельных структурных свойств.

Третий недостаток является продолжением второго и указывает на необходимость разработки принципиально новых способов представления структур, которые в большей мере были бы ориентированы на решение задач описания и сравнения структур.

В исследованиях по анализу свойств структур значительные успехи достигнуты, в частности, в области анализа структур химических соединений [3]. Разработано большое множество количе-

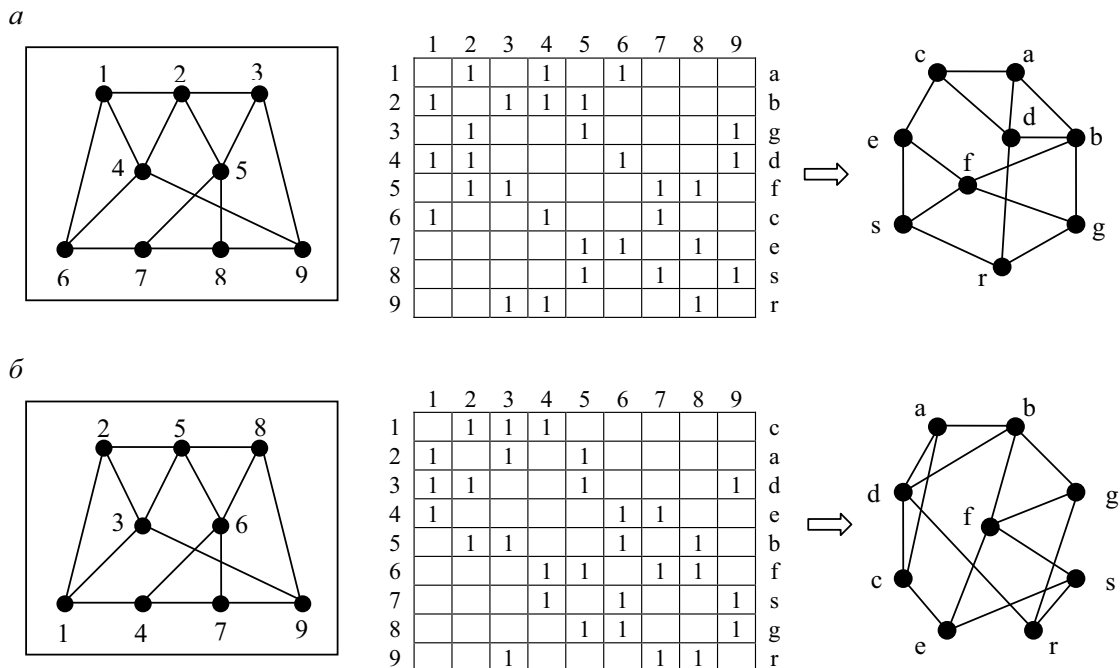


Рис. 1. Варианты описания схемы с разной нумерацией вершин

ственных показателей (топологических и молекулярных дескрипторов), используемых для описания структурных свойств графов и сравнения по ним структур химических веществ. Получены также некоторые результаты и по более сложной проблеме — построение на основе дескрипторов математических моделей (в частности регрессионных), связывающих характеристики структур химических соединений с их физико-химическими свойствами. Вместе с тем использование дескрипторов даёт возможность сравнивать структуры лишь опосредованно, т. к. каждый дескриптор определяет некоторую частную характеристику структуры. По дескрипторам нельзя восстановить структуру аналогично тому, как фоторобот, полученный по отдельным элементам, не является портретом конкретного лица.

Вернёмся к упомянутой выше идее получения для каждой вершины уникальной структурной характеристики и предположим, что такие описатели для вершин найдены и обозначены буквами ( $a, b, c, d, e, f, g, r, s$ ). На рис. 1 справа от матриц для каждой вершины указан соответствующий описатель и приведены графы, у которых вместо номеров вершин проставлены соответствующие описатели. Оба графа идентичны (отличаются расположением вершин) и представляют граф, который можно рассматривать как инвариант для описания схемы с любой нумерацией контактов.

На интуитивном уровне понятно, что, например, описатель  $g$  для вершины 3 на рис. 1,  $a$  (вершина 8 на рис. 1,  $b$ ) должен отражать её уникальное положение относительно всех других вершин графа. Понятно также, что уникальность положения вершины может быть обусловлена наличием в графе некоторых структурных различий в отношениях между его вершинами и уникальным расположением данной вершины относительно этих различий. Так, например, вершины 2 и 3 (рис. 1,  $a$ ) имеют разные степени, а вершины 6 и 3 имеют одинаковые степени, но различаются тем, что вершина 3 инцидентна двум вершинам 2 и 5, имеющим степень 4, а вершина 6 инцидентна только одной вершине 4, со степенью 4.

Если удастся объединить (интегрировать в описатель) все имеющиеся в графе структурные различия относительно каждой вершины, то можно надеяться на получение уникальной характеристики в виде интегрального описателя вершины. Совокупность таких описателей для всех вершин графа будем именовать интегральным описателем структуры — Integral Structure Descriptor (ISD). Разработке метода интеграции структурных различий и получению ISD для структур определённого класса посвящена настоящая статья.

Разумеется, что полезные применения ISD не ограничиваются возможностью получения описания структуры графа, инвариантного относительно нумерации вершин. Правда, получаемый при этом вывод о том, что два графа с равными ISD являются изоморфными, уже можно считать важным результатом и достаточным основанием для проведения исследований по методам получения

ISD. Заметим также, что вопросы применения ISD для анализа структур будут рассмотрены в последующих публикациях.

### Структурные различия и методы их интеграции

Предварительно уточним понятие структурного различия. Структура графовой модели является композицией из множества элементов (вершин) и отношений между ними (рёбер). Вершинам и рёбрам могут присваиваться различные атрибуты. Наличие атрибутов в общем случае приводит к структурным различиям между вершинами. Такие структурные различия будем именовать *атрибутными*. Очевидно, что наличие атрибутных различий упрощает задачу получения ISD при анализе соответствующих структур. При отсутствии атрибутов графовая модель является обыкновенным графом. Для такого графа получение ISD существенно усложняется, т. к. все вершины и рёбра становятся неразличимы и структурные различия, если они есть, должны быть обнаружены непосредственно в структуре отношений между вершинами.

Ряд структурных различий, способных привести к дифференциации вершин, соответствуют числовым характеристикам, определяемым на графах, в том числе и характеристикам некоторых топологических дескрипторов [3]. Простейшей такой характеристикой, например, являются степени вершин. Несколько сложнее вычисляется характеристика, разделяющая вершины на центральные и периферийные [4]. Так, для графа на рис. 1,  $a$ , вершины 2, 4, 5, 7, 9 являются центральными, а вершины 1, 3, 6, 8 — периферийными. Теория графов предлагает ряд других вычисляемых характеристик, приводящих к разбиению (дифференциации) вершин на группы.

Все структурные различия на основе вычисляемых характеристик, способных привести к дифференциации вершин, назовём *базовыми*. Структурные различия, основанные на структуре отношений между вершинами, в том числе и с учётом атрибутных и базовых различий, будем именовать *относительными*. Например, в графе на рис. 1,  $a$ , вершина 9 отличается от вершины 8 тем, что вершина 9 непосредственно связана с вершиной 3, которая в свою очередь связана с двумя четырёхстепенными вершинами 2 и 5, а вершина 8 с вершиной 3 связана через вершину 9. Заметим, что относительные структурные различия могут быть обнаружены и учтены лишь через механизм интеграции различий.

Предлагаемый метод интеграции различий для получения ISD может использовать атрибутные, базовые и относительные различия. Важным является то обстоятельство, что метод может работать при отсутствии атрибутных и базовых различий, т. е. осуществляет обнаружение и интеграцию не вычисляемых (скрытых) относительных различий. Изложение метода будем вести на примере получения ISD для обыкновенного графа, представленного матрицей смежности вершин на рис. 1,  $a$ . Результат применения метода свободной интеграции для данного примера приведён на рис. 2.

Обыкновенный граф, содержащий  $n$  вершин, будем представлять матрицей смежности вершин  $A=\|a_{ij}\|_{n \times n}$ . Элемент матрицы  $a_{ij}=1$ , если между вершинами  $i$  и  $j$  имеется ребро,  $a_{ij}=0$ , в противном случае. Матрица является симметричной, элементы  $a_{ii}=0$ . При отсутствии атрибутов все вершины графа принимаются неразличимыми и их исходные характеристики описываются вектором  $\alpha^0=\{\alpha_i^0\}$ , все элементы которого  $\alpha_i^0=1$ .

Процесс интеграции различий начинает работу с вектора  $\alpha$  и последовательно с помощью рекуррентной процедуры выполняет шаги интеграции по преобразованию вектора  $\alpha$  в  $\alpha^{f+1}$ . В итоге формируется последовательность векторов  $\alpha^0, \alpha^1, \dots, \alpha^f, \dots, \alpha^n$ .

Рассмотрим содержание шага интеграции по получению  $\alpha^{f+1}$  на основе  $\alpha^f$ . Вектор  $\alpha^f$  содержит элементы  $\alpha_i^f$ , которые являются кодовыми числами из ряда  $1, 2, \dots, n$ , где  $n$  – число вершин в графе. В общем случае значения элементов  $\alpha_i^f$  в векторе  $\alpha^f$  могут повторяться. Все элементы вектора  $\alpha^f$ , для которых  $\alpha_i^f=\text{const}$ , назовём *однородной группой*. Так, например, в векторе  $\alpha^0$  (рис. 2) все элементы  $\alpha_i^0=1$ , т. е. имеет место одна однородная группа, а вектор  $\alpha^2$  содержит 3 однородные группы.

Элемент  $\alpha_i^f$ , значение которого в векторе  $\alpha^f$  встречается один раз, соответствует уникальной характеристике  $i$ -й вершины и является её интегральным описателем. Если все элементы вектора  $\alpha^f, f \neq 0$ , являются уникальными характеристиками, то это означает, что для графа получен интегральный описатель структуры – ISD. Для нашего примера – это вектор  $\alpha^7$ .

Для выполнения шага интеграции используем операцию попарного произведения элементов двух векторов  $A_i=(a_{i1}, a_{i2}, \dots, a_{in})$  и  $\alpha^f=(\alpha_1^f, \alpha_2^f, \dots, \alpha_n^f)$ . Операцию обозначим символом  $\otimes$ , тогда имеем:

$$A \otimes \alpha = B(\alpha); B(\alpha) = \{B_i(\alpha)\}; B_i(\alpha) = A_i \otimes \alpha;$$

$$B_i(\alpha) = (\alpha_{i1}\alpha_1^f, \alpha_{i2}\alpha_2^f, \dots, \alpha_{in}\alpha_n^f); B_i(\alpha) = > B_j(\alpha).$$

Преобразование  $B_i(\alpha) = > B_j(\alpha)$  выполняется для удобства сравнения векторов  $B_i(\alpha)$ . Согласно преобразованию вектор  $B_i(\alpha)$  заменяется на множество  $B_i(\alpha)$  – совокупность упорядоченных по возрастанию значений ненулевых элементов вектора  $B_i(\alpha)$ . Принимается, что два вектора  $B_i(\alpha)$  и  $B_j(\alpha)$  равны, если равны их множества  $B_i(\alpha)$  и  $B_j(\alpha)$ , т. е.

$$\forall i \neq j [B_i(\alpha) = B_j(\alpha)] = > (B_i(\alpha) = B_j(\alpha)).$$

Заметим, что на рис. 2 элементы в множествах  $B_i(\alpha)$  не упорядочены, например,  $B_7(\alpha) = (6, 3, 5)$ . Разделители между элементами в множествах  $B_i(\alpha)$  не указаны, т. к. все элементы – одноразрядные числа.

Вторая часть шага интеграции выполняет присвоение множествам  $B_i(\alpha)$  характеристик (кодовых чисел)  $\alpha_i^{f+1}$ . С этой целью в каждой однородной группе элементов вектора  $\alpha^f$ , содержащей более одного элемента с равными кодовыми числами, выбирается первый элемент  $\alpha_i^f$  и соответствующему множеству  $B_i(\alpha)$  присваивается кодовое число  $\alpha_i^{f+1} = \alpha_i^f$ . Для каждого другого  $k$ -го элемента данной группы множеству  $B_k(\alpha)$  присваивается кодовое число  $\alpha_k^{f+1}$  согласно логическому выражению:

$$\forall k \in I(\alpha_i^f) \& k \neq i [B_k(\alpha) = B_i(\alpha)] = > (\alpha_k^{f+1} = \alpha_i^f), \text{ иначе}$$

$$(\alpha_k^{f+1} = \max[\{\alpha_i^f\} \cup \{\tilde{\alpha}_i^{f+1}\}] + 1).$$

Здесь  $\{\alpha_i^f\}$  – множество кодовых чисел в векторе  $\alpha^f$ , а  $\{\tilde{\alpha}_i^{f+1}\}$  – множество кодовых чисел, имеющих в векторе  $\alpha^{f+1}$ , которые присвоены ранее в данном шаге интеграции  $I(\alpha_i^f)$  – индексное множество элементов однородной группы вектора  $\alpha^f$ , имеющих кодовые числа равные  $\alpha_i^f$ .

На примере (рис. 2) проследим присвоение кодовых чисел множествам  $B_i(\alpha^1)$ . Вектор  $\alpha^1$  имеет 2 однородные группы с  $I(1) = (1, 3, 6, 7, 8, 9)$  и  $I(2) = (2, 4, 5)$ . Множеству  $B_1(\alpha^1) = (2, 2, 1)$  присваивается кодовое число  $\alpha_1^2 = \alpha_1^1 = 1$  как первому элементу группы. Второму элементу этой группы  $\alpha_3^1$  соответствует множество  $B_3(\alpha^1) = (2, 2, 1)$ , которое совпадает с  $B_1(\alpha^1)$  и, следовательно, ему присваивается кодовое число  $\alpha_3^2 = 1$ . Остальные множества  $B_6(\alpha^1)$ ,  $B_7(\alpha^1)$ ,  $B_8(\alpha^1)$ ,  $B_9(\alpha^1)$  равны между собой и не равны  $B_1(\alpha^1)$  и  $B_3(\alpha^1)$  и, следовательно, им присваиваются коды  $\alpha_6^2 = \alpha_7^2 = \alpha_8^2 = \alpha_9^2 = 3$ . Первому множеству  $B_2(\alpha^1) = (1, 1, 2, 2)$  второй однородной группы присваивается код  $\alpha_2^2 = \alpha_4^1 = 2$ . Второе множество этой группы  $B_4(\alpha^1)$  не совпадает с  $B_2(\alpha^1)$  и получает код  $\alpha_4^2 = 4$ . Третье множество  $B_5(\alpha^1)$  совпадает с  $B_4(\alpha^1)$  и также получает код  $\alpha_5^2 = 4$ . Кодовое число  $\alpha_2^2 = 2$  в векторе  $\alpha^2$  не повторяется, т. е. соответствует уникальной характеристике, которая принимается в качестве интегрального описателя вершины 2. Это означает, что кодовое число данной вершины в последующих шагах процесса свободной интеграции не изменится и будет включено в состав ISD.

Из рис. 2 следует, что для получения ISD потребовалось 7 шагов интеграции структурных различий. При этом все вершины графа получили уникальные характеристики о чём свидетельствует вектор  $\alpha^7$ , в котором нет совпадающих кодовых чисел, т. е. нет однородных групп. Вектор  $\alpha^7$ , элементы  $\alpha_i^7$  которого дополнены множествами

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | $\alpha^0$ | $\alpha^1$ |   | $\alpha^2$ |   | $\alpha^3$ |   | $\alpha^4$ |   | $\alpha^5$ |   | $\alpha^6$ |   | $\alpha^7$ |   |           |
|---|---|---|---|---|---|---|---|---|---|------------|------------|---|------------|---|------------|---|------------|---|------------|---|------------|---|------------|---|-----------|
| 1 |   | 1 |   | 1 |   | 1 |   |   |   | 1          | 111        | 1 | 221        | 1 | 243        | 1 | 243        | 1 | 243        | 1 | 243        | 1 | 243        | 1 | (2,4,3)   |
| 2 | 1 |   | 1 | 1 | 1 |   |   |   |   | 1          | 1111       | 2 | 1122       | 2 | 1144       | 2 | 1144       | 2 | 1146       | 2 | 1746       | 2 | 1746       | 2 | (1,7,4,6) |
| 3 |   | 1 |   |   |   | 1 |   |   | 1 | 1          | 111        | 1 | 221        | 1 | 243        | 1 | 243        | 1 | 263        | 7 | 263        | 7 | 268        | 7 | (2,6,8)   |
| 4 | 1 | 1 |   |   |   |   | 1 |   | 1 | 1          | 1111       | 2 | 1211       | 4 | 1233       | 4 | 1233       | 4 | 1233       | 4 | 1233       | 4 | 1238       | 4 | (1,2,3,8) |
| 5 |   | 1 | 1 |   |   |   |   | 1 | 1 | 1          | 1111       | 2 | 2111       | 4 | 2133       | 4 | 2155       | 6 | 2155       | 6 | 2755       | 6 | 2755       | 6 | (2,7,5,9) |
| 6 | 1 |   |   | 1 |   |   | 1 |   |   | 1          | 111        | 1 | 121        | 3 | 143        | 3 | 145        | 3 | 145        | 3 | 145        | 3 | 145        | 3 | (1,4,5)   |
| 7 |   |   |   |   | 1 | 1 |   | 1 |   | 1          | 111        | 1 | 211        | 3 | 433        | 5 | 435        | 5 | 635        | 5 | 635        | 5 | 635        | 5 | (6,3,9)   |
| 8 |   |   |   |   | 1 |   | 1 |   | 1 | 1          | 111        | 1 | 211        | 3 | 433        | 5 | 453        | 5 | 653        | 5 | 653        | 5 | 658        | 9 | (6,5,8)   |
| 9 |   |   | 1 | 1 |   |   |   | 1 |   | 1          | 111        | 1 | 121        | 3 | 143        | 3 | 145        | 3 | 145        | 3 | 745        | 8 | 745        | 8 | (7,4,9)   |

Рис. 2. Пример работы метода свободной интеграции структурных различий

$(A_i \otimes \alpha^7) \Rightarrow B_i(\alpha^7)$  (на рис. 2 указаны в скобках), принимается в качестве ISD графа. В таком виде ISD полностью представляет структуру графа и необходимость в матрице смежности вершин отпадает.

### Область интеграции различий для класса структур графовых моделей

Изложенный выше метод можно рассматривать как процесс свободной интеграции различий в структуре автономного графа, выполняемой независимо от других графов. Термин «свободной интеграции» в первую очередь отражает определённую свободу при назначении кодовых чисел множествам  $B_i(\alpha)$ . Вместе с тем, исходя из целей получения ISD и его использования при анализе класса структур, получение ISD для любой структуры из заданного класса должно быть согласовано.

Стратегия согласованного получения ISD в заданном классе структур основывается на следующих положениях:

- вводится понятие класса структур графов;
- даётся определение области интеграции структурных различий для одной структуры и класса структур;
- метод интеграции структурных различий преобразуется для согласованной работы в классе структур.

Введение понятия класса структур обусловлено тем, что проблема описания и сравнения структур на основе ISD представляет интерес при исследовании и сопоставлении структур, близких по определённым признакам, т. е. принадлежащих определённому классу. Для обыкновенных графов в качестве такого признака могут быть приняты значения степеней вершин. В соответствие с этим выделим класс структур графов, у которых степени вершин составляют совокупность значений, совпадающих с элементами определённого вектора  $\alpha^i$ . Например, для графа, представленного на рис. 2, вектор  $\alpha^i$  содержит 3 вершины со степенью 4 и 6 вершин со степенью 3. Тогда любой обыкновенный граф  $G$  с данной комбинацией значений степеней в векторе  $\alpha^i$  будет относиться к классу структур  $G(\alpha^i)$ .

Класс  $G(\alpha^i)$  можно существенно расширить, если вместо вектора  $\alpha^i$  указывать совокупность входящих в него несовпадающих элементов  $\{\alpha_i^1\} = \{1; 2\}$ . В этом случае класс  $G\{\alpha_i^1\} = G\{1; 2\}$  будет отличаться от  $G(\alpha^i)$  тем, что графы  $G \in G\{\alpha_i^1\}$  могут содержать в векторе  $\alpha^i$  конечное число вершин с произвольным соотношением чисел степеней 3 и 4. Графы, состоящие из вершин с равными степенями, воспринимаются как однородные группы. Интеграция различий в однородных группах осуществляется с помощью введения в вектор  $\alpha^i$  абстрактной характеристики [5] и будет рассмотрена в отдельном исследовании.

Область интеграции для каждой вершины определим как дерево, содержащее описатели вершин и связи между ними на всех шагах интеграции. Корнем дерева является интегральный описатель вершины. В примере на рис. 2 интегральный опи-

сатель для вершины 8 получен на 7-м шаге интеграции. Если принять, что все вершины графа имеют степень 3, то число внешних вершин дерева области интеграции вершины 8 составит  $3^7 = 2187$ , а для вершин всего графа  $2187 \times 9 = 19683$ . При совмещении одинаковых вершин область интеграции различий структуры графа можно представить в компактном виде. Такое представление области для рассматриваемого примера показано на рис. 3. В нём опущены вершины нулевого шага интеграции, а также исключены дублирующие описатели вершин. Жирными линиями выделены интегральные описатели вершин. Дублирующие описатели возникают в ситуации, когда интегральный описатель получен, но об этом становится известно через несколько шагов интеграции. Например, описатель  $\alpha_1^3$  с множеством  $B_i(\alpha^3) = (2, 4, 3)$  получен на 3-м шаге (рис. 2), но о том, что этот описатель оказался интегральным, определилось на 5-м шаге.

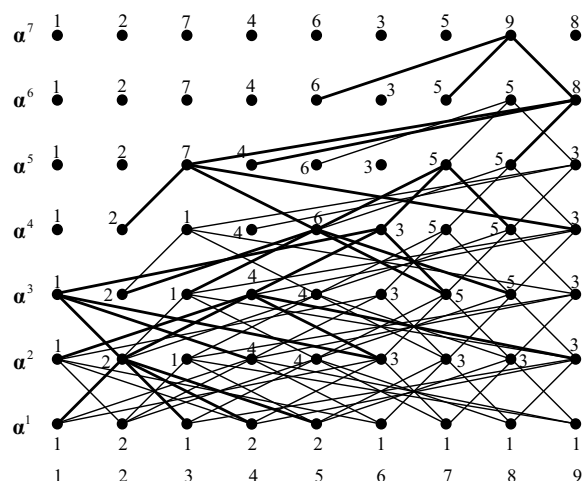


Рис. 3. Пример области интеграции различий для графа на рис. 2

Представление области интеграции в таком виде позволяет проследить процесс появления структурных различий и дифференциации вершин на однородные группы с последующей дифференциацией вершин в группах до получения уникальных характеристик. Работа процесса интеграции устроена так, что любые различия, фиксируемые кодовыми числами одновременно для всех вершин, становятся доступными в виде новых различий на следующем шаге интеграции. Таким образом, на каждом шаге интеграции можно наблюдать эффект «всасывания» различий, их «расползания» в соответствии с особенностями структуры и «продвижения» на более высокий шаг интеграции. Поэтому, несмотря на то, что описатель каждой вершины определяется всегда только на основе описателей смежных вершин, на самом деле он объединяет информацию, собранную в результате многократного «посещения» описателей всех других вершин графа. Например, ранее отмечалось, что относительно вершины 8 на 7-м шаге интеграции было учтено более 2000 таких посещений.

1:=(1, 1, 1) 1; (1, 1, 1, 1) 2;  
 2:=(1, 2, 2) 1; (1, 1, 2, 2) 2; (1, 1, 2, 3) 3; (1, 1, 2, 4) 4;  
 3:=(2, 3, 4) 1; (1, 1, 4, 4) 2; (1, 3, 4) 3; (1, 2, 3, 3) 4; (3, 3, 4) 5; (1, 1, 2, 3) 6;  
 (1, 2, 3) 7; (3, 4, 4) 8; (3, 3, 4, 4) 9; (2, 3, 3) 10;  
 4:=(1, 1, 4, 4) 2; (1, 4, 5) 3; (1, 2, 3, 3) 4; (3, 4, 5) 5; (1, 2, 5, 5) 6; (2, 5, 8, 8) 7;  
 (6, 8, 10) 8; (2, 3, 8, 8) 9; (5, 8, 9) 10; (6, 7, 10) 11; (3, 6, 6) 12;  
 (6, 6, 7, 10) 13; (3, 5, 9) 14; (6, 6, 7) 15;  
 5:=(1, 1, 4, 6) 2; (1, 2, 3, 3) 4; (3, 5, 6) 5; (1, 2, 5, 5) 6; (2, 3, 6) 7;  
 6:=(1, 4, 6, 7) 2; (1, 2, 3, 3) 4; (2, 5, 5, 7) 6; (2, 3, 6) 7; (4, 5, 7) 8;  
 7:=(1, 2, 3, 8) 4; (2, 6, 8) 7; (5, 6, 8) 9;

Рис. 4. Область интеграции для структур двух графов из класса  $G(\alpha)$

Форму представления области интеграции, показанную на рис. 3, следует рассматривать как модель, удобную для понимания сути процесса интеграции различий. Для согласованной работы метода интеграции в классе структур предложена другая, более удобная форма представления области интеграции и разработаны правила последовательного расширения границ области для заданного класса структур.

Анализ процесса интеграции показал, что состав описателей вершин в множествах  $B_i(\alpha)$  определяется составом однородных групп в векторе  $\alpha$ . Например (рис. 3), для множества  $B_7(\alpha)=(4,3,3)$  важно лишь то, что вершина 7 имеет одну связь с любой из вершин группы  $(\alpha_4^2, \alpha_5^2)=(4,4)$  и две связи с любыми двумя вершинами из группы  $(\alpha_6^2, \alpha_7^2, \alpha_8^2, \alpha_9^2)=(3,3,3,3)$ . Поэтому в области интеграции для каждой группы достаточно хранить описатель для одной (любой) вершины из этой группы. Совокупность таких описателей и интегральных описателей вершин, полученную на всех шагах свободной интеграции различий для одного из графов  $G$  класса  $G(\alpha)$ , задаёт область интеграции в начальных границах. Форма представления такой области для рассматриваемого примера приведена на рис. 4. Слева указаны номера шагов и для каждого  $f$ -го шага приведена совокупность множеств  $B(f)$  и назначенных кодовых чисел  $\alpha^{f-1}$ . Жирным шрифтом на рис. 4 выделена область в начальных границах, полученная на основе области, приведённой на рис. 3. В данной форме представления каждая запись в виде  $(B_k(f-1))\alpha_k^f$  соответствует тому, что на  $f$ -м шаге интеграции описатель  $\alpha_k^f$  был назначен всем вершинам однородной группы с множеством  $B_k(f-1)$ , либо оказался интегральным описателем, т. е. вошёл в состав ISD.

Работа метода интеграции различий в классе  $G(\alpha)$  отличается тем, что при назначении кодовых чисел  $\alpha_k^f$  соответствующее множество независимо от его принадлежности к однородной группе срав-

нивается с множествами  $B_k(f-1) \in B(f-1)$   $f$ -го шага области интеграции. В этом случае назначение кодовых чисел не является свободным, оно полностью зависит от области интеграции и производится согласно логическому выражению:

$$\forall B_k(\alpha^{f-1}) \in B(\alpha^{f-1}) \& \forall B_k(f-1) \in B(f-1) [B_k(\alpha^{f-1}) = B_k(f-1)] = > (\alpha_k^f = \alpha_k^f), \text{ иначе}$$

$$(\alpha_k^f = \max_k [\{\alpha_k^f\} \cup \{\tilde{\alpha}_k^f\}] + 1),$$

где  $\{\alpha_k^f\}$  – совокупность описателей в области интеграции на  $f$ -м шаге;  $\{\tilde{\alpha}_k^f\}$  – множество кодовых чисел уже назначенных в векторе  $\alpha$ .

Пример работы метода зависимой интеграции структурных различий для графа в классе  $G(\alpha)$  приведён на рис. 5. Как видно из данного примера, описатели на шаге 2 полностью совпадают с соответствующими описателями области интеграции. На шаге 3 появилось 5 новых описателей, а на 4-м шаге – 9 таких описателей, которые включены в область интеграции (рис. 4) и тем самым расширили её границы. Расширение области интеграции за счёт новых описателей будет происходить всякий раз, когда очередной граф из класса  $G(\alpha)$  не окажется изоморфным одному из графов, участвовавших ранее в формировании области. Из этого следует, что чем большее число графов в классе  $G(\alpha)$  пройдут через процедуру интеграции различий, тем более полно область интеграции будет представлять данный класс.

#### Некоторые сведения об ISD-технологии анализа структур

Сравнивая графы на рис. 2 и 5 по полученным ISD нетрудно видеть, что эти графы неизоморфны. О степени близости их структур можно судить по сходству описателей на соответствующих шагах

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | $\alpha^0$ | $\alpha^1$ | $\alpha^2$ | $\alpha^3$ | $\alpha^4$ |               |
|---|---|---|---|---|---|---|---|---|---|------------|------------|------------|------------|------------|---------------|
| 1 |   |   |   |   | 1 | 1 | 1 |   | 1 | 1   1111   | 2   1121   | 4   3121   | 6   5828   | 7          | (11,12,13,15) |
| 2 |   |   | 1 |   |   | 1 |   | 1 |   | 1   111    | 1   211    | 3   413    | 3   68,10  | 8          | (9,12,14)     |
| 3 |   | 1 |   |   |   | 1 | 1 |   | 1 | 1   1111   | 2   1121   | 4   3121   | 6   3828   | 9          | (8,12,13,15)  |
| 4 |   |   |   |   | 1 |   | 1 |   | 1 | 1   111    | 1   121    | 3   321    | 7   598    | 10         | (11,13,15)    |
| 5 | 1 |   |   | 1 |   |   |   | 1 |   | 1   111    | 1   211    | 3   433    | 5   67,10  | 11         | (7,10,14)     |
| 6 | 1 | 1 | 1 |   |   |   |   |   |   | 1   111    | 1   221    | 1   434    | 8   636    | 12         | (7,8,9)       |
| 7 | 1 |   | 1 | 1 |   |   |   | 1 |   | 1   1111   | 2   2211   | 2   4433   | 9   667,10 | 13         | (7,9,10,14)   |
| 8 |   | 1 |   |   | 1 |   | 1 |   |   | 1   111    | 1   112    | 3   332    | 10   359   | 14         | (8,11,13)     |
| 9 | 1 |   | 1 | 1 |   |   |   |   |   | 1   111    | 1   221    | 1   443    | 8   667    | 15         | (7,9,10)      |

Рис. 5. Пример работы метода зависимой интеграции в классе структур графов  $G(\alpha)$

интеграции. Проблема оценки близости структур на основе ISD — это предмет отдельного исследования. Здесь лишь отметим, что при сравнении структур графов из одного класса, область интеграции выступает в роли интегрального описателя структур этого класса, т. е. в роли некоторого образа, относительно которого выполняется интеграция различий в данных графах. В этом случае результаты интеграции становятся сопоставимы, их можно сравнивать между собой и на этой основе оценивать близость структур рассматриваемых графов.

Таким образом, предлагаемая концепция создания и использования области интеграции в качестве образа (эталона) класса структур, определяет содержание нового подхода к проведению анализа подобия структур на основе сравнения их интегральных описателей — ISD. Соответствующую технологию анализа структур будем именовать ISD-технологией. Анализ подобия структур согласно ISD-технологии включает выполнение следующих действий:

- выбирается совокупность признаков, по которым определяется класс структур графов (например, вектор  $\alpha^i$ );
- для любого представителя класса выполняется свободная интеграция структурных различий и формируется область интеграции в начальных границах (базовая часть области);
- для двух сравниваемых структур из данного класса последовательно выполняется зависимая интеграция различий и получение ISD. После выполнения очередной зависимой интеграции принимается решение о расширении границ области за счёт новых описателей вершин;
- на основе полученных ISD оценивается подобие и определяются другие свойства структур графов.

Приведённые этапы ISD-технологии отражают лишь стандартный сценарий действий, выполняемых при анализе структур. Конкретные требования и условия анализа могут вносить изменения в состав и содержание задач, реализованных в ISD-технологии. При этом следует иметь в виду, что разные пользователи для одного класса могут создавать разные области интеграции (эталоны), т. к. текущее состояние области зависит от последовательности рассмотрения графов. Это является вполне допустимым и объясняется тем, что результаты анализа двух графов, полученные относительно разных эталонов будут идентичны, т. к. в любом случае один из графов согласно ISD-технологии

успевает выступить в качестве составной части эталона для другого.

### Заключение

В статье изложены основные положения методов свободной и зависимой интеграции структурных различий и получения ISD для графов в заданном классе структур. В качестве описателя класса структур предлагается область интеграции, которая формируется в процессе анализа структур. При анализе структур она используется как эталон, относительно которого производится сравнение графов в заданном классе структур. Данные методы составляют ядро общей технологии анализа структур, названной ISD-технологией.

При изложении методов многие частные вопросы были опущены. Среди них наиболее существенными являются учёт атрибутивных и базовых структурных различий в процессе интеграции, условия возникновения устойчивых однородных групп и правила дифференциации вершин в таких группах, особенности использования ISD при наличии устойчивых однородных групп, условия включения новых описателей вершин в область интеграции различий. Все эти вопросы будут рассмотрены в последующих публикациях. Поэтому на данном этапе исследований преждевременно судить об эффективности применения ISD-технологии для анализа широкого класса структур. Основные опасения здесь сводятся к тому, что с расширением класса структур область интеграции может стать недопустимо большой. В этом случае потребуются разработка дополнительных правил формирования области в допустимых границах.

Что касается непосредственно самих методов свободной и зависимой интеграции структурных различий, то они, несомненно, являются эффективными. Из этого следует, что ISD-технология в большей степени ориентирована на получение оценок подобия структур и выявление труднообнаруживаемых («тонких») свойств в структурах графов из узкого класса. В связи с этим интерес может представлять задача классификации графов широкого класса на подклассы с допустимыми областями интеграции. Решение данной задачи позволит проводить анализ структур графов в широком классе, предварительно разбив его на более мелкие (узкие) подклассы.

*Работа выполнена при проведении НИР, в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» 2009 – 2013 гг. Госконтракт № П 2396.*

### СПИСОК ЛИТЕРАТУРЫ

1. Погребной В.К. Об одном методе определения изоморфизма графов // Кибернетика. — 1982. — № 2. — С. 7–13.
2. Погребной В.К. О декомпозиции графов на классы изоморфных подграфов // В сб.: Вопросы программирования и автоматизации проектирования. — Томск, ТГУ, 1979. — № 4. — С. 82–96.
3. Дэмер М., Эммерт-Штрайб Ф., Цой Ю. Р., Вармуза К. Новый функционал информативности для анализа структуры химиче-

ских графов // Известия Томского политехнического университета. — 2010. — Т. 316. — № 5. — С. 5–11.

4. Кристофидес Н. Теория графов. Алгоритмический подход. — М.: Мир, 1978. — 432 с.
5. Погребной В.К. Алгоритм решения задачи определения изоморфизма гиперграфов // Известия Томского политехнического университета. — 2010. — Т. 316. — № 5. — С. 16–21.

*Поступила 01.04.2011 г.*